Visualizing Data Driven Phenotypes

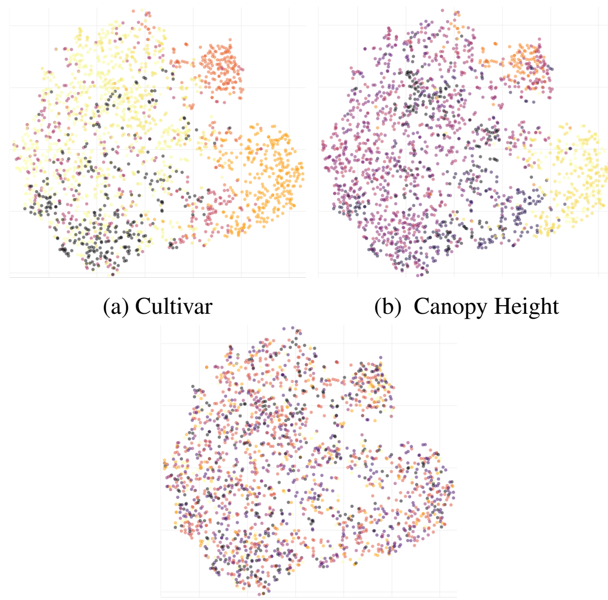
Zeyu Zhang, Hong Xuan, Xiaotong Liu, Abby Stylianou and Robert Pless
George Washington University
{zeyu|xuanhong|liuxiaotong2017|astylianou|pless}@gwu.edu

Deep metric learning uses large datasets to train deep convolutional neural networks to map images from the same class to similar locations in feature space, and images from different classes to distant locations. Given sufficient variety in the training data, these networks have the capacity to learn representations that are robust to variations in viewpoint, lighting, and other imaging conditions. By training such a network to differentiate between different plant cultivars from large scale plant monitoring datasets, the network must learn the visual and structural similarities and differences between the different varieties. This extended abstract discusses initial approaches to training and visualizing such a model.

We train a Resnet-50 [1] network, using 141,818 depth and reflectance images (concatenated into false color [depth, depth, reflectance] images) of sorghum from the TERRA-REF project [2]. The data is labeled by which of 1,597 plots the imagery came from, where the data from a plot consists of many images of a single cultivar from June 1st to June 5th, 2017 (there are multiple plots per cultivar). We use triplet loss to learn a 32-dimensional embedding where images from each plot are embedded nearby. t-SNE [3] is used to project this embedding into 2-dimensional space to visualize how the network has learned to represent the data.

Figure 1a shows this 2 dimensional projection of the learned embedding for a sampling of the training data that includes between 1 and 3 plots for each of 10 cultivars. The unimodal cultivar clusters (e.g., the red cluster in the upper right and orange cluster in the bottom right), show the network is generalizing plot labels to learn visual and structural representations consistent all data from the cultivar. Figure 1b shows the same points color coded by hand-measured canopy height suggesting this may be one feature automatically learned by the embedding. Figure 1c is color-coded by date of data capture; the fact that the points are not clustered by date indicates that the network has learned to ignore daily imaging variations.

We believe this approach can be extended to find subtle variation in phenotypes and to characterize plant development over time.



(a) Cultivar (b) Canopy Height

(c) Date

Figure 1: t-SNE visualization of data subset color coded by different labels. For cultivar, each color corresponds to a unique cultivar. For canopy height, each color corresponds to a different range of canopy heights. For date, each color corresponds to a particular day. There is no correspondence between colors on the different plots.

References

[1] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[2] Maxwell Burnette et al. TERRA-REF Data Processing Infrastructure. In *Proceedings of the Practice and Experience on Advanced Research Computing*, PEARC '18, pages 27:1–27:7. ACM, 2018.

[3] L. van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, 9:2579–2605, 2008.